Structure
Article

Secondary Structure of Huntingtin Amino-Terminal Region

Mee Whi Kim,[1,*] Yogarany Chelliah,[2] Sang Woo Kim,[2] Zbyszek Otwinowski,[1] and Ilya Bezprozvanny[3]
[1]Department of Biochemistry
[3]Department of Physiology
[2]Howard Hughes Medical Institute
University of Texas Southwestern Medical Center at Dallas, Dallas, TX 75390, USA
*Correspondence: MeeWhi.Kim@UTSouthwestern.edu
DOI 10.1016/j.str.2009.08.002

SUMMARY

Huntington's disease is a genetic neurodegenerative disorder resulting from polyglutamine (polyQ) expansion (>36Q) within the first exon of Huntingtin (Htt) protein. We applied X-ray crystallography to determine the secondary structure of the first exon (EX1) of Htt17Q. The structure of Htt17Q-EX1 consists of an amino-terminal α helix, poly17Q region, and polyproline helix formed by the proline-rich region. The poly17Q region adopts multiple conformations in the structure, including α helix, random coil, and extended loop. The conformation of the poly17Q region is influenced by the conformation of neighboring protein regions, demonstrating the importance of the native protein context. We propose that the conformational flexibility of the polyQ region observed in our structure is a common characteristic of many amyloidogenic proteins. We further propose that the pathogenic polyQ expansion in the Htt protein increases the length of the random coil, which promotes aggregation and facilitates abnormal interactions with other proteins in cells.

INTRODUCTION

Huntington's disease (HD) is an autosomal-dominant neurodegenerative disorder. Neuropathological analysis of HD patients reveals selective and progressive neuronal loss in the striatum (Vonsattel and DiFiglia, 1998), particularly affecting the GABAergic medium spiny striatal neurons (MSN). At the molecular level, the cause of HD is a polyglutamine (polyQ) expansion ($\geq$36Q) near the amino terminus of Huntingtin (Htt), a 350 kDa ubiquitously expressed cytoplasmic protein (The Huntington's Disease Collaborative Research Group, 1993). The cellular mechanisms that link Huntingtin polyQ expansion (Htt^{exp}) with the disease are under intense investigation. Most evidence is consistent with the hypothesis that polyQ-expanded Htt acquires a "toxic gain of function" (Tobin and Signer, 2000). A number of toxic functions have been assigned to Htt^{exp}, including formation of toxic aggregates, effect on gene transcription, induction of apoptosis, and disruption of key neuronal functions such as proteosomal function, ubiquitination, axonal transport, endocytosis, synaptic transmission, and Ca^{2+} signaling (Bezprozvanny, 2009; Cha, 2007; Li and Li, 2004; Ross, 2002; Rubinsztein, 2002; Tobin and Signer, 2000; Truant et al., 2008). Many of the proposed mechanisms suggest that the mutant Htt^{exp} is involved in pathological interactions with other signaling proteins in cells, leading to neuronal dysfunction and death.

Information about the structure of the polyQ region of Htt is critical for understanding Htt^{exp} toxicity, and might aid in the development of potential HD therapies. However, the structure determination of the polyQ region of Htt has proven to be an extremely difficult problem (Temussi et al., 2003). In the aggregated form, polyQ sequence most likely adopts β sheet structure (Bevivino and Loll, 2001; Chen et al., 2001; Perutz, 1996; Perutz et al., 2002; Singer and Dewji, 2006; Takahashi et al., 2008; Tanaka et al., 2001). The biophysical studies of soluble monomeric form of polyQ fragments of various length provided evidence for random coil conformation (Altschuler et al., 1997; Bennett et al., 2002; Chen et al., 2001; Masino et al., 2002), for α-helical conformation (Bhattacharyya et al., 2006; Nagai et al., 2007), β sheet (Nagai et al., 2007), and extended helix (Chellgren et al., 2006). Random coil, α and β sheet, α helix, μ helix, and π helix conformations have been predicted using computational methods (Armen et al., 2005; Khare et al., 2005; Lathrop et al., 1998; Tsukamoto et al., 2006; Wang et al., 2006). Recently the crystal structure of the synthetic peptide $GQ_{10}G$ in the extended conformation was determined in the complex with a monoclonal antibody (Li et al., 2007). Although these studies provided important insights into potential conformations of the polyQ sequence, most of them have been performed in the absence of a native protein context or yielded low-resolution data. Here we present the secondary structure of the first exon (EX1) of Htt, containing 17 glutamines (Htt17Q) determined by X-ray crystallography at atomic resolution. We discovered that the poly17Q region adopts multiple conformations in the structure, including α helix, random coil, and extended loop. We conclude that the conformation of the poly17Q region is influenced by the conformation of neighboring protein regions, demonstrating the importance of the native protein context. We propose that the conformational flexibility of the polyQ region observed in our structure is a common characteristic of many amyloidogenic proteins.

RESULTS

Crystallization of Htt17Q Exon 1 Fragment

The first exon of Huntingtin protein containing 17 glutamines (Htt17Q-EX1) was expressed in bacteria and crystallized as

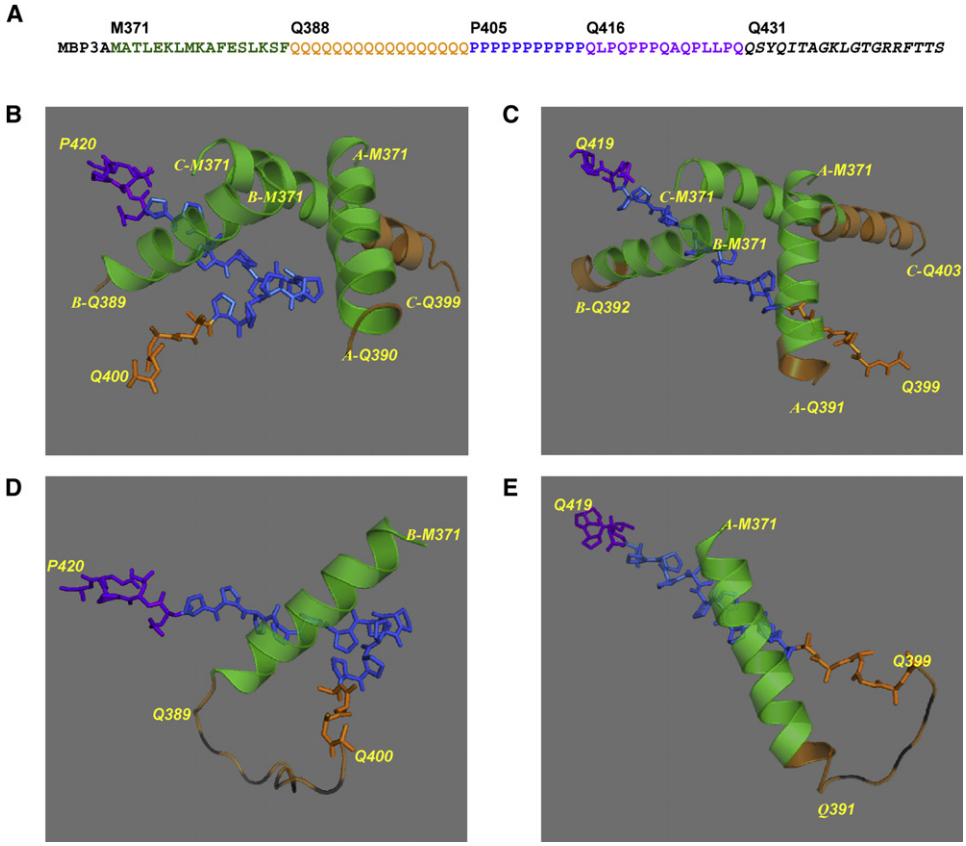

Figure 1. Secondary Structure of Htt17Q-EX1

(A) Amino acid sequence of MBP-Htt17Q-EX1. MBP3A denotes the maltose binding protein followed by a 3Ala linker. M371 to Q430 is the sequence of Htt17Q-EX1, which is subdivided into a 17 aa N-terminal region (M371 to F387), poly17Q region (Q388 to Q404), poly11P region (P405 to P415), and 15 aa mixed P/Q region (Q416 to Q430). The sequence from Q431 to the C terminus is the 19 aa tag added to facilitate crystallization.

(B) The structure of Htt17Q-EX1 trimer from c^{95} crystal. The structures of MBP and 3A linker are removed for clarity. The amino-terminal α helix of Htt17Q-EX1 extends from Met371 to Phe387 (Green). The following poly17Q region (orange) is α helical and unstructured (random coil). The poly11P region (blue) adopts a PP helix (shown as stick model) in the kinked conformation. The initial part of the polyP/Q region (purple) is also in PP-helix conformation (shown as stick model). The terminal part of the poly17Q region (orange) is in the extended conformation (shown as stick model).

(C) The structure of Htt17Q-EX1 trimer from c_{Hg}^{99} crystal. Same as in (B) but the poly11P region is in the straight conformation.

(D) The complete structure of B molecule of Htt17Q-EX1 monomer from c^{95} crystal. The striped orange loops are for the random coil region between Gln389 and Gln399, which is invisible on the map.

(E) The complete structure of A molecule of Htt17Q-EX1 monomer from c_{Hg}^{99} crystal. The striped orange loops are for the random coil region between Gln391 and Gln398, which is invisible on the map.

maltose-binding protein (MBP) fusion protein (see Experimental Procedures for details). The amino-terminal MBP molecule was connected with Htt17Q-EX1 via a three-alanine (3A) linker. The amino acid (aa) numbering in MBP-Htt17Q-EX1 construct starts with MBP, so that Met1 in the Htt17Q corresponds to Met371 in the MBP-Htt17Q-EX1. The sequence of Htt17Q-EX1 can be subdivided into several distinct regions (Figure 1A). The N-terminal region of Htt17Q-EX1 starting at Met371 (Met1 in Htt sequence) extends for 17 aa until Phe387. The next region Gln388–Gln404 contains a stretch of 17Gln (poly17Q), which is immediately followed by a stretch of 11Pro from Pro405 until Pro415 (poly11P). The proline stretch is followed by 15 aa mixed Pro and Gln region (polyP/Q) that spans from Gln416 to Gln430. The native Htt sequence ends at Gln430 and is followed by 19 aa C-terminal tag that was added to facilitate crystallization (Figure 1A).

X-ray diffraction data were collected for 30 crystals of MBP-Htt17Q-EX1. The diffraction for all crystals was anisotropic, extending to 2.8 Å in the high-resolution directions [h, 0, 0] and [0, k, 0] and 4.5 Å in the low-resolution direction [0, 0, l]. All crystals were monoclinic, space group C121, with slightly different unit cell size and β angles ranging from 90° to 99°. The asymmetric unit contained three molecules of MBP-Htt17Q-EX1 (designated as molecules A, B, C) in all crystals. The phasing was done by molecular replacement (MR) using MBP structural models (see Experimental Procedures for details). MR gave solutions for seven crystals with complete maps for MBP and partial maps for Htt17Q-EX1. The β angles and other crystallographic parameters for these seven crystals are listed in Table 1. Following MR, the models of Htt17Q-EX1 fragments in these seven crystals were built separately but simultaneously. The final refinement statistics for the seven crystals used in Htt17Q-EX1

Table 1. Diffraction Data Collection and Refinement Statistics for Seven Crystals Used for Htt17Q-EX1 Structure Determination

Crystal	c^{90}	c_a^{92}	c_b^{92}	c^{94}
Data Collection				
Wavelength (λ)	1.55	0.98	0.98	1.55
Space group	C121	C121	C121	C121
Unit cell parameter (Å, °)				
a	162	163	163	160
b	101	101	101	100
c	142	138	137	140
β	90	92	92	94
Resolution limit (Å)	3.8-41 / 3.8-4.1	4.0-37 / 4.0-4.6	3.5-37 / 3.5- 3.7	3.7–40 / 3.7-3.8
R_{sym}(%)	12.8 / 24.1	6.3 / 12.8	21.2 / 30.1	12.4 / 47.1
I/σI (overall/outer shell)	13.2 / 3.1	13.2 / 5.9	8.4 / 2.3	14.3 / 3.1
Redundancy (overall/outer shell)	2.0 / 1.9	1.8 / 1.7	3.2 / 2.8	2.4 / 1.7
Overall completeness (%)	64	92	69	83
Refinement				
Overall mean B-factor (Å^2)	104	81	56	70
No. of data	16,075	21,188	18,671	19,565
R_{work} / R_{free}	0.251/0.303	0.245/ 0.292	0.247 / 0.295	0.272/0.298
RMS_D (Å)	0.009	0.006	0.009	0.007
RMS_A (°)	1.180	0.937	1.149	1.171
PDB ID	3IO4	3IO6	3IOT	3IOU
Crystal	c^{95}	c^{99}	c_{Hg}^{99}	
Data Collection				
Wavelength (λ)	0.98	0.98	0.98	
Space group	C121	C121	C121	
Unit cell parameter (Å, °)				
a	163	163	101	
b	100	101	162	
c	137	134	139	
β	95	99	99	
Resolution limit (Å)	3.5-37/3.5-3.6	3.5-43/3.5-3.6	3.5-40/3.5-3.7	
R_{sym}(%)	9.9/33.2	8.6/35.0	16.0/24.8	
I/σI (overall /outer shell)	9.7/2.9	10.0/1.3	15.0/4.8	
Redundancy(overall/outer shell)	2.7/1.9	2.8/1.5	3.7/3.2	
Overall completeness (%)	82	87	92	
Refinement				
Overall mean B-factor (Å^2)	56	82	48	
No. of data	20472	20629	24515	
R_{work} / R_{free}	0.257 / 0.267	0.263 / 0.280	0.243 / 0.280	
RMS_D (Å)	0.007	0.007	0.006	
RMS_A (°)	1.335	0.979	0.961	
PDB ID	3IOR	3IOV	3IOW	

The crystals are sorted by β angle of unit cell (C121). The values of β angles are assigned to each crystal name in superscript. For the same β angles crystals are noted by letters of the alphabet in subscript. The mercury socked crystal is indicated by Hg in subscript. PDB accession numbers for the data collected with each crystal are shown in the last row.

structure determination are summarized in Table 1. The Protein Data Bank (PDB) accession numbers for each of the seven data sets are also included in Table 1. The final seven structures of Htt17Q-EX1 were very similar despite differences in β angles in the corresponding crystals.

Secondary Structure of the Htt17Q-EX1 Fragment

MBP-Htt17Q-EX1 forms a trimer, composed of molecules A, B, and C, and packed sandwich-like in the crystal (see Figure S1 available online). The trimer is likely to be an artifact of crystallization because MBP-Htt17Q-EX1 is a monomer in solution

Table 2. The Last Visible Residue in Poly17Q α Helix and Conformation of Poly11P in Different MBP-Htt17Q-EX1 Crystals

Crystal	c^{90}	c_a^{92}	c_b^{92}	c^{94}	c^{95}	c^{99}	c_{Hg}^{99}
β angle	90	92	92	94	95	99	99
α helix A	395	388	388	388	388	390	389
α helix B	397	401	386	390	388	396	392
α helix C	395	396	391	400	396	393	402
PP helix	Straight	[a]	Straight	Kinked	Kinked	Kinked	Straight

[a] Not resolved on the map.

according to dynamic light scattering and size exclusion chromatography analysis (data not shown). Two representative models of Htt17Q-EX1 trimer are shown in Figures 1B (crystal c^{95}) and 1C (crystal c_{Hg}^{99}). The models consist of the N-terminal region of Htt17Q-EX1 starting with Met371 (green), the poly17Q region (orange), and stretches of poly11P (blue) and a mixed polyP/Q region (purple). The N-terminal region of Htt17Q-EX1 (green) is resolved as an α helix for all three molecules (A, B, C) in the asymmetric unit. The structures of N-terminal α-helical trimers are identical in all seven crystals and can be superimposed on each other. The initial portions of poly17Q region (orange) on Figures 1B and 1C can be identified as an α helix or loop for all three molecules. The rest of these poly17Q regions are not visible on the map and not shown on Figures 1B and 1C. Only a single poly11P region (blue) and an initial fragment of polyP/Q region (purple) are resolved on the map in PP helix form and the same region is invisible for other two molecules in the asymmetric unit. The structure of poly11P in crystals c^{94}, c^{95}, and c^{99} is kinked in the middle (Figure 1B, Table 2). In crystals c^{90}, c_b^{92}, and c_{Hg}^{99} the poly11P is straight (Figure 1C, Table 2). In crystal c_a^{92} the poly11P region adopts a mixture of "kinked" and "straight" conformations and can not be clearly resolved on the map (Table 2). The last 4 glutamine residues from the poly17Q region immediately preceding poly11P reappear on the map in the extended conformation (orange).

Examples of A and B molecules of Htt17Q-EX1 are shown on Figures 1D and 1E. The B molecule from crystal c^{95} is shown on Figure 1D. The structure of this molecule includes N-terminal α-helix (green), polyQ stretch (orange), "kinked" poly11P region (blue), and initial part of polyP/Q region (purple). The unstructured random coil portion of polyQ region (Gln389–Gln399) is shown on Figure 1D as a striped orange loop. The remaining carboxy-terminal polyQ region (Gln400–Gln404) is resolved in the extended loop conformation (orange stick model). The A molecule from c_{Hg}^{99} is shown on Figure 1E. This structure includes N-terminal α helix (green), polyQ stretch (orange), straight poly11P region (blue), and initial portion of the polyP/Q region (purple). The beginning of the polyQ region (Gln388–Gln390) is α helical and the carboxy-terminal portion of polyQ region (Gln399–Gln404) is in the extended loop conformation (orange stick model). The random coil portion of polyQ region in this molecule (Gln391–Gln398) is shown by the striped orange loop.

Conformation of the Polyglutamine Region

For all three molecules (A, B, and C) in all seven crystals, the N-terminal portion of Htt17Q-EX1 (green) is α helical (Figures 1B and 1C). In all molecules, the N-terminal α helix continues until Phe387 but eventually loses helical conformation within the polyQ region (orange). The remainder of polyQ region following the α helices forms a "random coil" conformation. The polyQ region is connected to the poly11P region in kinked (e.g., Figure 1D) or straight (e.g., Figure 1E) PP-helical conformations. The last 4 Gln residues at the carboxy-terminal end of poly17Q region (Gln400–Gln404) are in the extended conformation (Figures 1D and 1E). The transition from α helix to a random coil within poly17Q region occurs at variable positions for the six molecules shown in Figure 1B and 1C. The examples of three α helices with various lengths chosen from three different crystals and corresponding electron density maps are presented on Figure 2. The shortest helix is in the A molecule in crystal c^{95}. The α helix in this crystal continues only until Gln388, the first residue of the poly17Q region, and then becomes a loop (Figure 2A). The diminishing electron density map of the polyQ loop of the short α helix on Figure 2A approaches hydrophilic residues Asp41, Lys42, Glu45, and Gln49 in MBP-B or MBP-A molecules. The medium-length α helix is C molecule of crystal c^{90}, which extends until Gln395 and then becomes a loop (Figure 2B). The longest α helix is observed in the C molecule of crystal c_{Hg}^{99}; it extends until Gln402 and then becomes random coil that is invisible on the map (Figure 2C). Table 2 indicates the last residues of the poly17Q α helixes for molecules A, B, and C in each of the solved seven crystals with different β angles.

Conformation of the Polyproline Region

The poly11P region is resolved as PP helix in the crystals. Only single poly11P helix is resolved in the trimer, which adopts a kinked or straight conformation in different crystals (Table 2) The beginning of poly11P in the kinked conformation is positioned at the equal distance between A and B N-terminal α helices (Figure 1B), and the beginning of poly11P in the straight conformation is positioned between A and C N-terminal α helices (Figure 1C). The poly11P region is connected with the N-terminal α helices via the polyQ random coil as described above. Following poly11P region, a few residues from the polyQ/P mixed region (shown in purple on Figures 1B and 1C) maintain an extended conformation until the structure is no longer visible. The parts of poly11P helix are stabilized by interactions with hydrophobic surface of the α-helical trimers in the crystals (Figures 1B and 1C).

DISCUSSION

The two parallel bar diagrams on Figure 3 show the protein sequence of MBP-Htt17Q-EX1 and the corresponding secondary structure of each sequence region as determined from our crystallographic data. The N-terminal region of Htt (M371

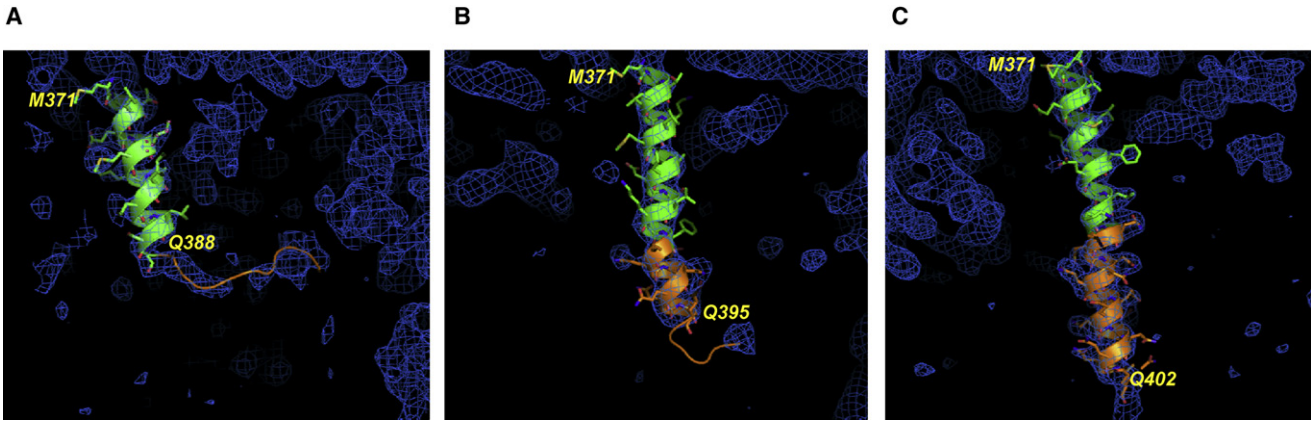

Figure 2. Structure of Amino-Terminal and Poly17Q Regions of Htt17Q-EX1

The structures of the amino-terminal region (green) and poly17Q region (orange) of Htt17Q-EX1 are shown for three different molecules with variable lengths of poly17Q α helix. Also shown are corresponding regions of electron density maps contoured at 1.0 σ (blue).

(A) The short poly17Q helix makes a transition to loop at Gln388 (molecule A, crystal c^{95}).

(B) The medium poly17Q helix makes a transition to loop at Gln395 (molecule C, crystal c^{90}).

(C) The long poly17Q helix extends for the length of the polyQ region until Gln402 (molecule C, crystal c_{Hg}^{99}).

to F387) is α helical in all crystals (bold green). The N-terminal α helix ends at variable residues within poly17Q region between Q388 and Q400 (shaded green) and the rest of poly17Q region continues as a random coil (not colored). The polyQ residues immediately preceding poly11P region (Q400 to Q404) adopt extended loop conformation (shaded blue). The poly11P region exists as PP helix (bold blue) in all crystals, either in the kinked or straight conformation. The secondary structure beyond poly11P region continues as PP helix for several residues into mixed P/Q region (shaded blue) until the structure can no longer be resolved.

It is informative to compare our results with previous experimental and modeling studies of Huntingtin and other polyQ-containing proteins. The molecular dynamics simulations of Htt amino-terminal headpiece predicted α-helical structure (Kelley et al., 2009) and the α-helical structure of Htt amino-terminal peptide was observed by circular dichroism (CD) spectra analysis (Atwal et al., 2007). However, predominantly random coil conformation for the same region was reported by another group based on CD and nuclear magnetic resonance spectra analysis (Thakur et al., 2009). The Htt N-terminal headpiece modulates subcellular localization and aggregation of Htt^{exp} (Atwal et al., 2007; Rockabrand et al., 2007; Thakur et al., 2009), and definitive information about its secondary structure is useful for understanding these effects. The N-terminal region of Htt is α helical in all of our crystals (Figure 3), consistent with the molecular dynamics simulations and biophysical measurements (Atwal et al., 2007; Kelley et al., 2009). The agreement with these studies indicates that ability to form an α helix is an intrinsic property of the Htt amino-terminal region and not a result of

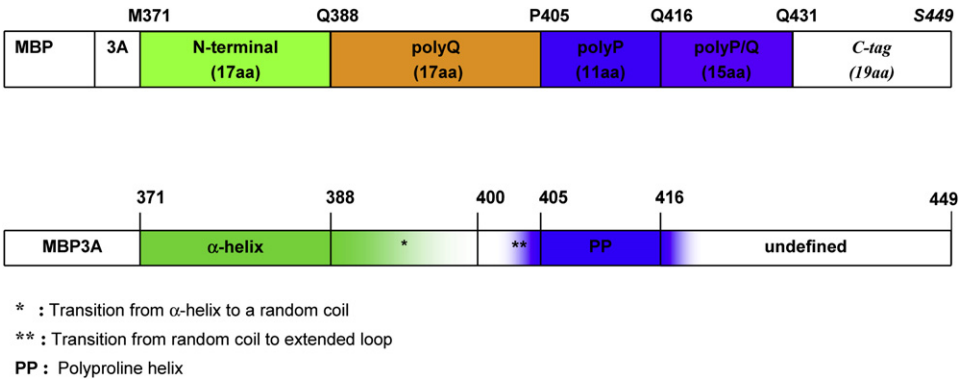

Figure 3. Schematic Diagram of the Secondary Structure Elements of Htt17Q-EX1

The first bar represents the sequence of MBP-Htt17Q-EX1 construct, which consists of MBP protein (not shown to scale), a three-alanine linker (3A), a 17-aa-long N-terminal region of Htt (green), a polyQ region with 17 glutamine residues (poly17Q, orange), a polyP region with 11 proline residues (poly11P, blue), a 15-aa-long mixed polyglutamine and polyproline region (polyP/Q, purple), and a C-terminal tag of 19 aa (not colored). The second bar is a summary of structural information obtained from analysis of Htt17Q-EX1 structures resolved in the seven crystals. The main secondary structure elements of Htt17Q-EX1 are shown: α helix (bold green), transition from α helix to random coil region (shaded green), random coil (uncolored), transition to extended loop region (shaded blue), PP helix (bold blue), transition from PP helix to unstructured region (shaded blue).

crystallization of Htt17Q-EX1 as amino-terminal MBP-fusion protein.

A range of structures was reported for polyQ region in previous experimental and modeling studies. Some studies indicated that monomeric polyQ peptide is unstructured and exists as a random coil in solution (Altschuler et al., 1997; Bennett et al., 2002; Chen et al., 2001; Masino et al., 2002). Using CD spectra analysis, other groups concluded that polyQ stretch is at least partially α helical (Bhattacharyya et al., 2006; Nagai et al., 2007), an observation supported by modeling results (Lathrop et al., 1998; Wang et al., 2006). We have found evidence for both random coil and α-helical conformations of polyQ in our structures (Figure 2), in agreement with both sets of observations. We further discovered that transition between α-helical and random coil conformations of polyQ region occurs at random positions (Figure 2, Table 2), indicating a very shallow equilibrium between these two polyQ conformations as suggested by molecular dynamics stimulations and biophysical measurements (Bhattacharyya et al., 2006; Wang et al., 2006). Additional biophysical and modeling studies provided evidence that short polyQ stretches might also exist as the extended conformation (Chellgren et al., 2006; Wang et al., 2006). The ability of polyQ fragments to adopt an extended conformation was further supported by crystal structure of $GQ_{10}G$ peptide bound to a monoclonal antibody (Li et al., 2007). In our structures the last 4 glutamine residues also adopt the extended conformation (Figure 3). The analysis of these structures led us to conclude that the extended conformation at the carboxy-terminal end of polyQ stretch is imposed by PP helix of the neighboring poly11P region. Similar conclusions were reached in biophysical studies of polyQ-polyP peptides in solution (Darnell et al., 2007). An ability of poly11P region to stabilize the structure of a polyQ stretch in the extended conformation might account for a known role played by poly11P region in reducing aggregation and toxicity of Htt^{exp} (Bhattacharyya et al., 2006; Darnell et al., 2007; Dehay and Bertolotti, 2006; Duennwald et al., 2006).

The polyQ in the monomeric state is likely to exist in an equilibrium among α-helical, random coil, and extended conformations, which is consistent with the range of structures observed in Htt17Q-EX1 crystals (Figure 3) and with results obtained in biophysical studies performed with the soluble proteins (Altschuler et al., 1997; Bennett et al., 2002; Bhattacharyya et al., 2006; Chellgren et al., 2006; Chen et al., 2001; Darnell et al., 2007; Masino et al., 2002; Nagai et al., 2007). A similar conclusion was reached in the recent conformational epitope mapping study performed with the mutant Htt-EX1 fragment in solution (Legleiter et al., 2009). The shallow equilibrium among α-helical, loop, and extended conformations could be easily affected by the length of polyQ sequence, neighboring protein context, interactions with other proteins, changes in temperature, or ionic composition. This "conformational flexibility" within Htt-polyQ region observed in our structure might be a common characteristic of many amyloidogenic proteins such as other polyQ-disease proteins and proteins involved in Alzheimer disease, Parkinson disease, and Prion disease (PrP) (Blondelle et al., 1997; Fraser et al., 1991; Murphy, 2002; Nguyen et al., 1995; Temussi et al., 2003). Similar to Htt-polyQ, the sequences of these proteins also have been predicted to have an ambiguous secondary structure and can easily adopt multiple conformations, some of which might be toxic (Williams and Paulson, 2008).

The polyQ expansions of more than 36Q are toxic to cells and cause HD and other neurodegenerative disorders (Gusella and MacDonald, 2000). One explanation for the toxicity of polyQ expansion is aggregation. The proteins containing expanded polyQ repeats are prone to aggregation and some studies suggested that polyQ aggregates are toxic to neurons (Ross, 2002; Truant et al., 2008). It is generally accepted that in the aggregated state polyQ-expanded proteins adopt β sheet structure (Bevivino and Loll, 2001; Chen et al., 2001; Perutz, 1996; Perutz et al., 2002; Singer and Dewji, 2006; Tanaka et al., 2001). Some modeling and biophysical studies suggested that monomeric polyQ can also adopt β sheet conformation in solution (Nagai et al., 2007; Wang et al., 2006). We have not observed an evidence of β strand conformation in the Htt17Q-EX1 crystals of wild-type Htt (Figure 3). However, it is possible that in the mutant Htt^{exp} long polyQ random coil might adopt different conformation, such as a β sheet (Nagai et al., 2007), whose nucleation can occur from the random coil state of expanded polyQ region (Uversky and Fink, 2004; Vitalis et al., 2008) or from some other conformation that needs to be determined experimentally. However, it has also been reported that pathogenic and nonpathogenic monomeric polyQ tracts have similar structural properties in solution (Klein et al., 2007), arguing against major conformational changes in long polyQ tracks. To address these questions, the crystal structure of Htt-EX1 containing pathological polyQ expansion needs to be determined.

Another explanation for the pathogenic effects of polyQ expansion is abnormal protein interactions. For example, toxicity of Htt^{exp} might result from pathological interactions between Htt^{exp} and signaling proteins in cells (Bezprozvanny, 2009; Cha, 2007; Li and Li, 2004; Ross, 2002; Rubinsztein, 2002; Tobin and Signer, 2000; Truant et al., 2008). From the structure we found that the α-helical form of polyQ is not involved in any strong molecular interactions. The random coil form of polyQ is localized to the solvent area or near the polar residues of MBP, suggesting weak interactions between polyQ random coil and the polar residues of MBP. From these results we conclude that the α-helical and random coil forms of Htt-polyQ region are relatively "inert." There are two potential mechanisms that might account for interactions between expanded polyQ region and other proteins. The first possibility is that the mutant polyQ region adopts a novel conformation, which makes it more "sticky." As discussed above, this "toxic" conformation can be a β sheet or some other conformation. The second possibility is that many weak protein interactions mediated by expanded polyQ in the random coil conformation are sufficient to result in pathological effects. In both cases the interactions between expanded polyQ region (in novel toxic conformation or in random coil conformations) would be stronger with certain amino acid sequences in target proteins, resulting in some degree of specificity for these interactions.

In conclusion, we applied X-ray crystallography to determine the secondary structure of the first exon of Huntingtin protein containing 17Q sequence. The structure of Htt17Q-EX1 consists of an amino-terminal α helix, a poly17Q region, and a PP helix formed by the poly11P and mixed P/Q regions. The poly11P region can adopt kinked or straight conformations. The poly17Q

region displays "conformational flexibility" by adopting α-helical, random coil, and extended loop conformations in our crystals. The conformations of poly17Q region appear to be influenced by the conformation of neighboring protein regions, demonstrating the importance of the native protein context. We propose that the conformational flexibility of the polyQ region observed in our structure is a common characteristic of many amyloidogenic proteins. We further propose that the pathogenic polyQ expansion in the Htt^{exp} protein increases the length of the random coil, which promotes Htt^{exp} aggregation and abnormal interactions with other proteins in cells. Similar ideas are likely to be relevant for understanding structure and toxicity of other polyQ-containing proteins and amyloidogenic proteins in general.

EXPERIMENTAL PROCEDURES

Cloning, Protein Expression, Purification, Crystallization and X-ray Diffraction Data Collection

The first exon (EX1) fragment of human *Huntingtin* containing 17 CAG repeats (Met1-Gln60) was amplified by polymerase chain reaction and cloned into NotI and PstI sites of the pMAL vector with modified 3Ala-linker (Center et al., 1998). The aa numbering in the resulting MBP-Htt17Q-EX1 fusion protein is continuous from the MBP protein, so that Met1 in Htt17Q sequence corresponds to Met371 and Gln60 in Htt17Q sequence corresponds to Gln430. Following Gln430 in the MBP-Htt17Q-EX1 sequence, the 19 aa carboxy-terminal (C-terminal) tag (QSYQITAGKLGTGRRFTTS) was added to facilitate crystallization. MBP-Htt17Q-EX1 protein was expressed in BL21 bacteria and purified on Maltose-binding affinity column, followed by molecular exclusion chromatography (Superdex™200, Pharmacia Biotech) using fast protein liquid chromatography. The purified MBP-Htt17Q-EX1 was crystallized in hanging drop without removing MBP in order to improve protein solubility. Single crystals 300/300/20 μm in size grew in the mother drop containing 12% polyethylene glycol, 200 mM Zn acetate, 200 mM sodium acetate, 100 mM sodium cacodylate (pH 6.5–7.4) at 4°C to 16°C in 2 to 6 weeks. Synthetic 10Gaunadines RNA was used as an additive in the molar ratio 1:1 to the protein during crystallization. The crystals were frozen in mother liquid in liquid nitrogen and the X-ray diffraction data were collected at the Advanced Photon Source 19ID beam line. Some crystals were soaked in 1 mM CH_3HgCl for 1 week prior to freezing. The diffraction data collected from 30 crystals were indexed and processed by HKL2000 (Minor et al., 2000).

Phasing by Molecular Replacement

All crystals of MBP-Htt17Q-EX1 had the monoclinic symmetry, space group C121, with three molecules of MBP-Htt17Q-EX1 in the asymmetric unit (ASU). For each crystal, phasing was done by MR using the Phaser program. Forty MBP PDB files were downloaded from the Protein Data Bank and different permutations of three MBP models were used as possible models for three proteins in the ASU. The solutions for seven data sets produced by MR were suitable to build a model of Htt17Q-EX1. The PDB files used in MR for these seven crystals are shown in Table S1. The MBP structures described in 1JW4, 1JW5, 10MP, 1LLS represent open forms without maltose and differ primarily in the side-chain conformation. The crystals were nearly isomorphous but insufficient for averaging. Thus, the structure determination was performed separately for each crystal. The electron density maps generated by MR consisted of rich and poor electron density layers, with the gradual transition from rich to poor electron density (Figure S1). The rich electron density layer corresponded to the layer of MBP molecules and the poor density layer corresponded to the layer of Htt17Q-EX1 in the crystal. The model for Htt17Q-EX1 was built using Coot. Table 1 summarizes the properties of the seven crystals and corresponding diffraction data sets used in solving Htt17Q-EX1 structure.

Model Building and Refinement

When the seven maps generated by MR were compared, it became apparent that the Htt17Q-EX1 fragments adopt somewhat different configuration in different crystals. To account for this fact, the models of Htt17Q-EX1 fragments in the seven crystals were built separately but simultaneously. Furthermore, to account for the weak diffraction from Htt17Q-EX1 fragment, the occupancy of a corresponding model was chosen in the range between 0.80 and 0.30 to yield the best R_{work} and R_{free} factors during refinement using refmac5. In the process of model building and refinement, the R_{work} and R_{free} were improved and the electron density map of Htt17Q-EX1 became more clear. Table 1 summarizes the final refinement statistics for the seven crystals used for Htt17Q-EX1 structure determination.

ACCESSION NUMBERS

Coordinates have been deposited in the Protein Data Bank under ID codes 3IO4 (C90), 3IO6(Ca92), 3IOT(Cb92), 3IOU(C94), 3IOR(C95), 3IOV(C99), and 3IOW(CHg99).

SUPPLEMENTAL DATA

Supplemental Data include one figure and one table and can be found with this article online at http://www.cell.com/structure/supplemental/S0969-2126(09)00299-8.

ACKNOWLEDGMENTS

We thank Johann Deisenhofer (J.D.) for generous support of this project and comments on the manuscript, Dominika Borek for help with data collection and processing, the personnel of ID-19 beamline at the Advanced Photon Source of Argonne National Laboratory, Bostjan Kobe for the gift of the MBP-3A vector, all members of the J.D. laboratory for help and useful advice, Kimberly Aikman for administrative assistance, and Youxing Jiang for comments on the paper. This work was supported by the Howard Hughes Medical Institute (J.D.), Robert A. Welch Foundation (I.B. and J.D.), and NIH grants NS056224 (I.B.) and GM053163 (Z.O.).

Received: May 28, 2009
Revised: July 23, 2009
Accepted: August 3, 2009
Published: September 8, 2009

REFERENCES

Altschuler, E.L., Hud, N.V., Mazrimas, J.A., and Rupp, B. (1997). Random coil conformation for extended polyglutamine stretches in aqueous soluble monomeric peptides. J. Pept. Res. *50*, 73–75.

Armen, R.S., Bernard, B.M., Day, R., Alonso, D.O., and Daggett, V. (2005). Characterization of a possible amyloidogenic precursor in glutamine-repeat neurodegenerative diseases. Proc. Natl. Acad. Sci. USA *102*, 13433–13438.

Atwal, R.S., Xia, J., Pinchev, D., Taylor, J., Epand, R.M., and Truant, R. (2007). Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity. Hum. Mol. Genet. *16*, 2600–2615.

Bennett, M.J., Huey-Tubman, K.E., Herr, A.B., West, A.P., Jr., Ross, S.A., and Bjorkman, P.J. (2002). Inaugural Article: A linear lattice model for polyglutamine in CAG-expansion diseases. Proc. Natl. Acad. Sci. USA *99*, 11634–11639.

Bevivino, A.E., and Loll, P.J. (2001). An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates formation of parallel beta -fibrils. Proc. Natl. Acad. Sci. USA *98*, 11955–11960.

Bezprozvanny, I. (2009). Calcium signaling and neurodegenerative diseases. Trends Mol. Med. *15*, 89–100.

Bhattacharyya, A., Thakur, A.K., Chellgren, V.M., Thiagarajan, G., Williams, A.D., Chellgren, B.W., Creamer, T.P., and Wetzel, R. (2006). Oligoproline effects on polyglutamine conformation and aggregation. J. Mol. Biol. *355*, 524–535.

Blondelle, S.E., Forood, B., Houghten, R.A., and Perez-Paya, E. (1997). Polyalanine-based peptides as models for self-associated beta-pleated-sheet complexes. Biochemistry *36*, 8393–8400.

Center, R.J., Kobe, B., Wilson, K.A., Teh, T., Howlett, G.J., Kemp, B.E., and Poumbourios, P. (1998). Crystallization of a trimeric human T cell leukemia virus type 1 gp21 ectodomain fragment as a chimera with maltose-binding protein. Protein Sci. *7*, 1612–1619.

Cha, J.H. (2007). Transcriptional signatures in Huntington's disease. Prog. Neurobiol. *83*, 228–248.

Chellgren, B.W., Miller, A.F., and Creamer, T.P. (2006). Evidence for polyproline II helical structure in short polyglutamine tracts. J. Mol. Biol. *361*, 362–371.

Chen, S., Berthelier, V., Yang, W., and Wetzel, R. (2001). Polyglutamine aggregation behavior in vitro supports a recruitment mechanism of cytotoxicity. J. Mol. Biol. *311*, 173–182.

Darnell, G., Orgel, J.P., Pahl, R., and Meredith, S.C. (2007). Flanking polyproline sequences inhibit beta-sheet structure in polyglutamine segments by inducing PPII-like helix structure. J. Mol. Biol. *374*, 688–704.

Dehay, B., and Bertolotti, A. (2006). Critical role of the proline-rich region in Huntingtin for aggregation and cytotoxicity in yeast. J. Biol. Chem. *281*, 35608–35615.

Duennwald, M.L., Jagadish, S., Muchowski, P.J., and Lindquist, S. (2006). Flanking sequences profoundly alter polyglutamine toxicity in yeast. Proc. Natl. Acad. Sci. USA *103*, 11045–11050.

Fraser, P.E., Nguyen, J.T., Surewicz, W.K., and Kirschner, D.A. (1991). pH-dependent structural transitions of Alzheimer amyloid peptides. Biophys. J. *60*, 1190–1201.

Gusella, J.F., and MacDonald, M.E. (2000). Molecular genetics: unmasking polyglutamine triggers in neurodegenerative disease. Nat. Rev. Neurosci. *1*, 109–115.

Kelley, N.W., Huang, X., Tam, S., Spiess, C., Frydman, J., and Pande, V.S. (2009). The predicted structure of the headpiece of the Huntingtin protein and its implications on Huntingtin aggregation. J. Mol. Biol. *388*, 919–927.

Khare, S.D., Ding, F., Gwanmesia, K.N., and Dokholyan, N.V. (2005). Molecular origin of polyglutamine aggregation in neurodegenerative diseases. PLoS Comput. Biol. *1*, 230–235.

Klein, F.A., Pastore, A., Masino, L., Zeder-Lutz, G., Nierengarten, H., Oulad-Abdelghani, M., Altschuh, D., Mandel, J.L., and Trottier, Y. (2007). Pathogenic and non-pathogenic polyglutamine tracts have similar structural properties: towards a length-dependent toxicity gradient. J. Mol. Biol. *371*, 235–244.

Lathrop, R.H., Casale, M., Tobias, D.J., Marsh, J.L., and Thompson, L.M. (1998). Modeling protein homopolymeric repeats: possible polyglutamine structural motifs for Huntington's disease. Proc. Int. Conf. Intell. Syst. Mol. Biol. *6*, 105–114.

Legleiter, J., Lotz, G.P., Miller, J., Ko, J., Ng, C., Williams, G.L., Finkbeiner, S., Patterson, P.H., and Muchowski, P.J. (2009). Monoclonal antibodies recognize distinct conformational epitopes formed by polyglutamine in a mutant huntingtin fragment. J. Biol. Chem. *284*, 21647–21658.

Li, S.H., and Li, X.J. (2004). Huntingtin-protein interactions and the pathogenesis of Huntington's disease. Trends Genet *20*, 146–154.

Li, P., Huey-Tubman, K.E., Gao, T., Li, X., West, A.P., Jr., Bennett, M.J., and Bjorkman, P.J. (2007). The structure of a polyQ-anti-polyQ complex reveals binding according to a linear lattice model. Nat. Struct. Mol. Biol. *14*, 381–387.

Masino, L., Kelly, G., Leonard, K., Trottier, Y., and Pastore, A. (2002). Solution structure of polyglutamine tracts in GST-polyglutamine fusion proteins. FEBS Lett. *513*, 267–272.

Minor, W., Tomchick, D., and Otwinowski, Z. (2000). Strategies for macromolecular synchrotron crystallography. Structure *8*, R105–110.

Murphy, R.M. (2002). Peptide aggregation in neurodegenerative disease. Annu. Rev. Biomed. Eng. *4*, 155–174.

Nagai, Y., Inui, T., Popiel, H.A., Fujikake, N., Hasegawa, K., Urade, Y., Goto, Y., Naiki, H., and Toda, T. (2007). A toxic monomeric conformer of the polyglutamine protein. Nat. Struct. Mol. Biol. *14*, 332–340.

Nguyen, J., Baldwin, M.A., Cohen, F.E., and Prusiner, S.B. (1995). Prion protein peptides induce alpha-helix to beta-sheet conformational transitions. Biochemistry *34*, 4186–4192.

Perutz, M.F. (1996). Glutamine repeats and inherited neurodegenerative diseases: molecular aspects. Curr. Opin. Struct. Biol. *6*, 848–858.

Perutz, M.F., Finch, J.T., Berriman, J., and Lesk, A. (2002). Amyloid fibers are water-filled nanotubes. Proc. Natl. Acad. Sci. USA *99*, 5591–5595.

Rockabrand, E., Slepko, N., Pantalone, A., Nukala, V.N., Kazantsev, A., Marsh, J.L., Sullivan, P.G., Steffan, J.S., Sensi, S.L., and Thompson, L.M. (2007). The first 17 amino acids of Huntingtin modulate its sub-cellular localization, aggregation and effects on calcium homeostasis. Hum. Mol. Genet. *16*, 61–77.

Ross, C.A. (2002). Polyglutamine pathogenesis: emergence of unifying mechanisms for Huntington's disease and related disorders. Neuron *35*, 819–822.

Rubinsztein, D.C. (2002). Lessons from animal models of Huntington's disease. Trends Genet. *18*, 202–209.

Singer, S.J., and Dewji, N.N. (2006). Evidence that Perutz's double-beta-stranded subunit structure for beta-amyloids also applies to their channel-forming structures in membranes. Proc. Natl. Acad. Sci. USA *103*, 1546–1550.

Takahashi, T., Kikuchi, S., Katada, S., Nagai, Y., Nishizawa, M., and Onodera, O. (2008). Soluble polyglutamine oligomers formed prior to inclusion body formation are cytotoxic. Hum. Mol. Genet. *17*, 345–356.

Tanaka, M., Morishima, I., Akagi, T., Hashikawa, T., and Nukina, N. (2001). Intra- and intermolecular beta-pleated sheet formation in glutamine-repeat inserted myoglobin as a model for polyglutamine diseases. J. Biol. Chem. *276*, 45470–45475.

Temussi, P.A., Masino, L., and Pastore, A. (2003). From Alzheimer to Huntington: why is a structural understanding so difficult? EMBO J. *22*, 355–361.

Thakur, A.K., Jayaraman, M., Mishra, R., Thakur, M., Chellgren, V.M., Byeon, I.J., Anjum, D.H., Kodali, R., Creamer, T.P., Conway, J.F., et al. (2009). Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. Nat. Struct. Mol. Biol. *16*, 380–389.

The Huntington's Disease Collaborative Research Group (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell *72*, 971–983.

Tobin, A.J., and Signer, E.R. (2000). Huntington's disease: the challenge for cell biologists. Trends Cell Biol. *10*, 531–536.

Truant, R., Atwal, R.S., Desmond, C., Munsie, L., and Tran, T. (2008). Huntington's disease: revisiting the aggregation hypothesis in polyglutamine neurodegenerative diseases. FEBS J. *275*, 4252–4262.

Tsukamoto, K., Shimizu, H., Ishida, T., Akiyama, Y., and Nukina, N. (2006). Aggregation mechanism of polyglutamine disease revealed using quantum chemical calculations, fragment molecular orbital calculations, molecular dynamics simulations, and binding free energy calculations. J. Mol. Struct. *778*, 85–95.

Uversky, V.N., and Fink, A.L. (2004). Conformational constraints for amyloid fibrillation: the importance of being unfolded. Biochim. Biophys. Acta *1698*, 131–153.

Vitalis, A., Wang, X., and Pappu, R.V. (2008). Atomistic simulations of the effects of polyglutamine chain length and solvent quality on conformational equilibria and spontaneous homodimerization. J. Mol. Biol. *384*, 279–297.

Vonsattel, J.P., and DiFiglia, M. (1998). Huntington disease. J. Neuropathol. Exp. Neurol. *57*, 369–384.

Wang, X., Vitalis, A., Wyczalkowski, M.A., and Pappu, R.V. (2006). Characterizing the conformational ensemble of monomeric polyglutamine. Proteins *63*, 297–311.

Williams, A.J., and Paulson, H.L. (2008). Polyglutamine neurodegeneration: protein misfolding revisited. Trends Neurosci. *31*, 521–528.

